

Chapter 19. Volume Organization and Naming

The *Volume Organization and Naming* Standard defines the standard way of organizing data sets onto physical media and the conventions for forming volume names and identifiers. A *volume* is one unit of a physical medium such as a CD, a DVD, or a magnetic tape. Data sets may reside on one or more volumes and multiple data sets may also be stored on a single volume. Volumes are grouped into *volume sets*.

Each volume has a directory structure containing subdirectories and files. Both random access (CD, DVD) and sequential access (magnetic tape) media are supported. A PDS volume on a sequential access medium has a virtual directory structure defined in the VOLUME object included in the file “VOLDESC.CAT”. This virtual structure may then be used to recreate the volume directory structure when the files are moved to a random access medium.

PDS recommends that the entire contents of an archive volume be based on a single version of the PDS Standards Reference. Software tools that work with one version of the Standards may not work with all versions.

19.1 Volume Set Types

Data may be organized into one of four types of archive volumes, based on the number of data sets on each volume and the number of volumes required to capture all the data. The directory organization of the volumes and the required files varies slightly depending on this volume type. Figures 19.1 through 19.5 depict the various volume directory structure options. The four volume types are described below.

1. *One data set on one volume.* This basic volume organization is illustrated in Figure 19.1. The required and optional files and directories are detailed in Section 19.3.
2. *One data set on many volumes.* In this case the INDEX subdirectory includes both local indices, for the data on the present volume, and cumulative indices, for the data on all (preceding) volumes. This layout is illustrated in Figure 19.2.
3. *Many data sets on one volume.* In this case, additional file naming conventions are imposed to prevent collisions; data subdirectories are organized by data set. There are two variations on this scheme:
 - a. *One logical volume* – That is, the data sets collected on the physical medium constitute a single logical volume and would generally be distributed together (Figure 19.3).
 - b. *Many logical volumes* – The physical medium contains several largely independent collections of data sets, with each collection organized as though it were on its own volume. This is useful when a larger capacity medium (say, DVD) is being used to hold several volumes produced on a smaller medium (e.g., CDROM). In this case, directories that are common to and identical on all

volumes need only be reproduced once (e.g., the SOFTWARE directory in Figure 19.4).

4. *Many data sets on many volumes.* This organization is most useful when several large data sets are being produced in parallel over an extended period of time (as with some space missions). Sections of each data set appear on each physical volume, requiring additional naming considerations.

Note that it is possible to have one or more volumes containing only data accompanied by an ancillary volume containing the DOCUMENT, CATALOG, GAZETTER, SOFTWARE, CALIB, and GEOMETRY directories relevant to all the other volumes. When this is done, the PDS requires that all files referenced by include-type pointers (see the *Pointer Usage* chapter in this document) be present on the data volume. The PDS recommends that ancillary files be archived on the same volume as the corresponding data wherever possible, to facilitate science access.

The contents and organization of the directories of all the volume types are described in the remainder of this chapter.

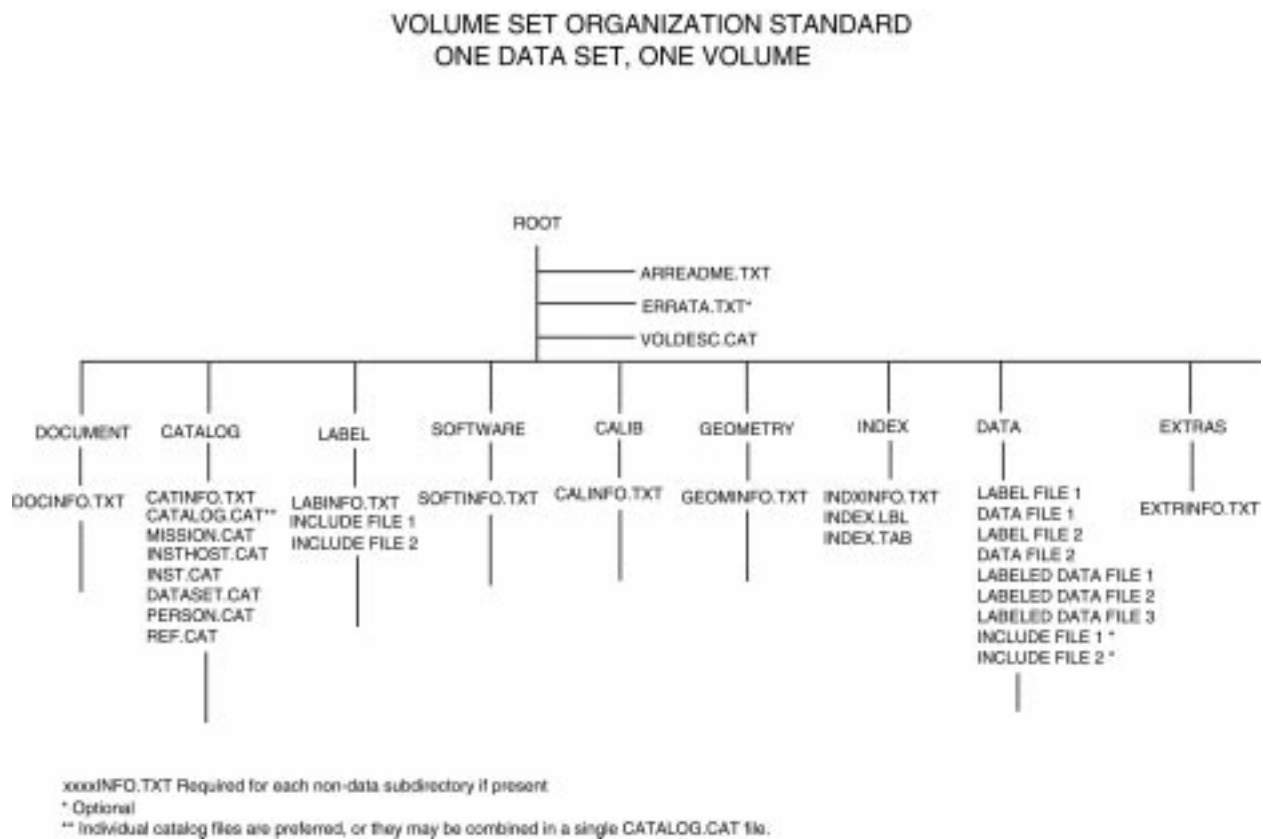


Figure 19.1 Volume Set Organization Standard - One Data Set, One Volume

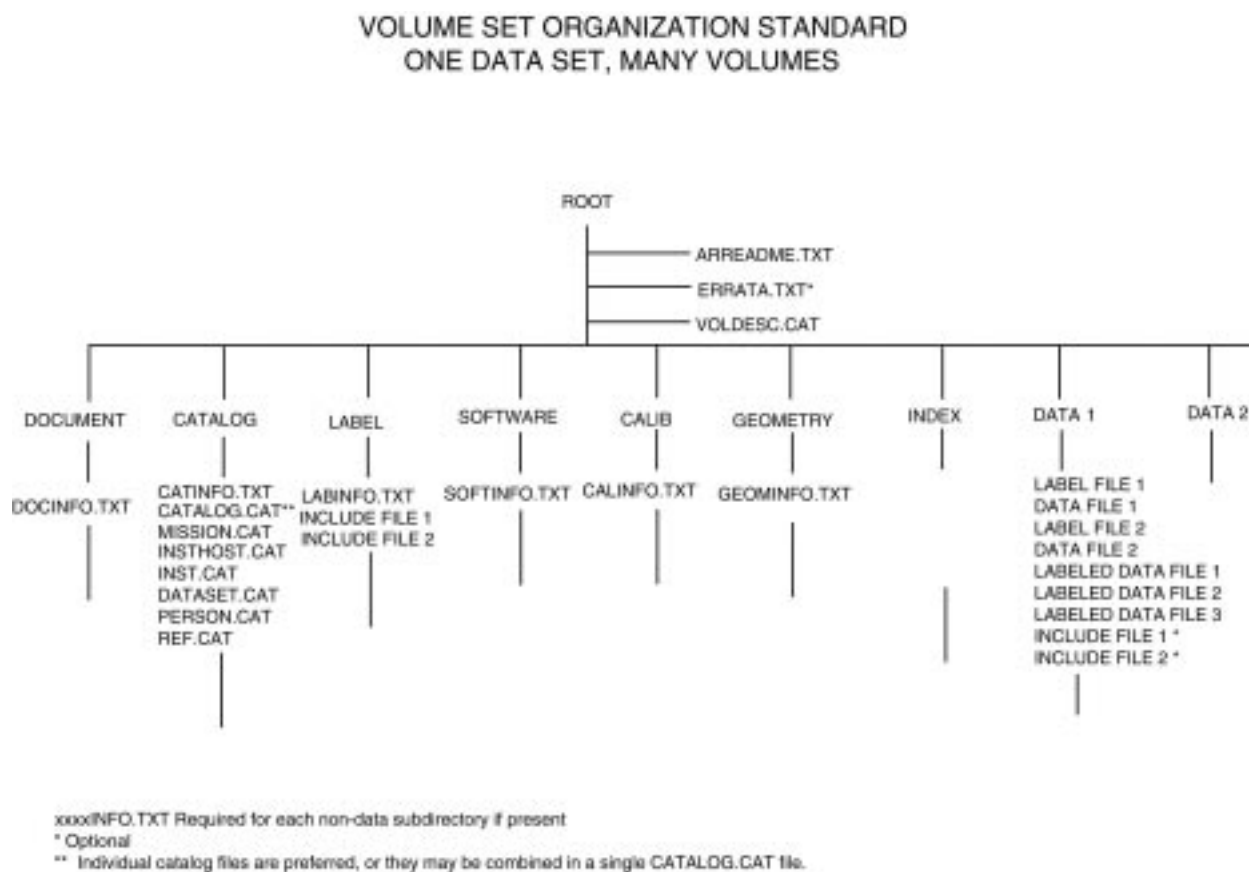


Figure 19.2 Volume Set Organization Standard - One Data Set, Many Volumes

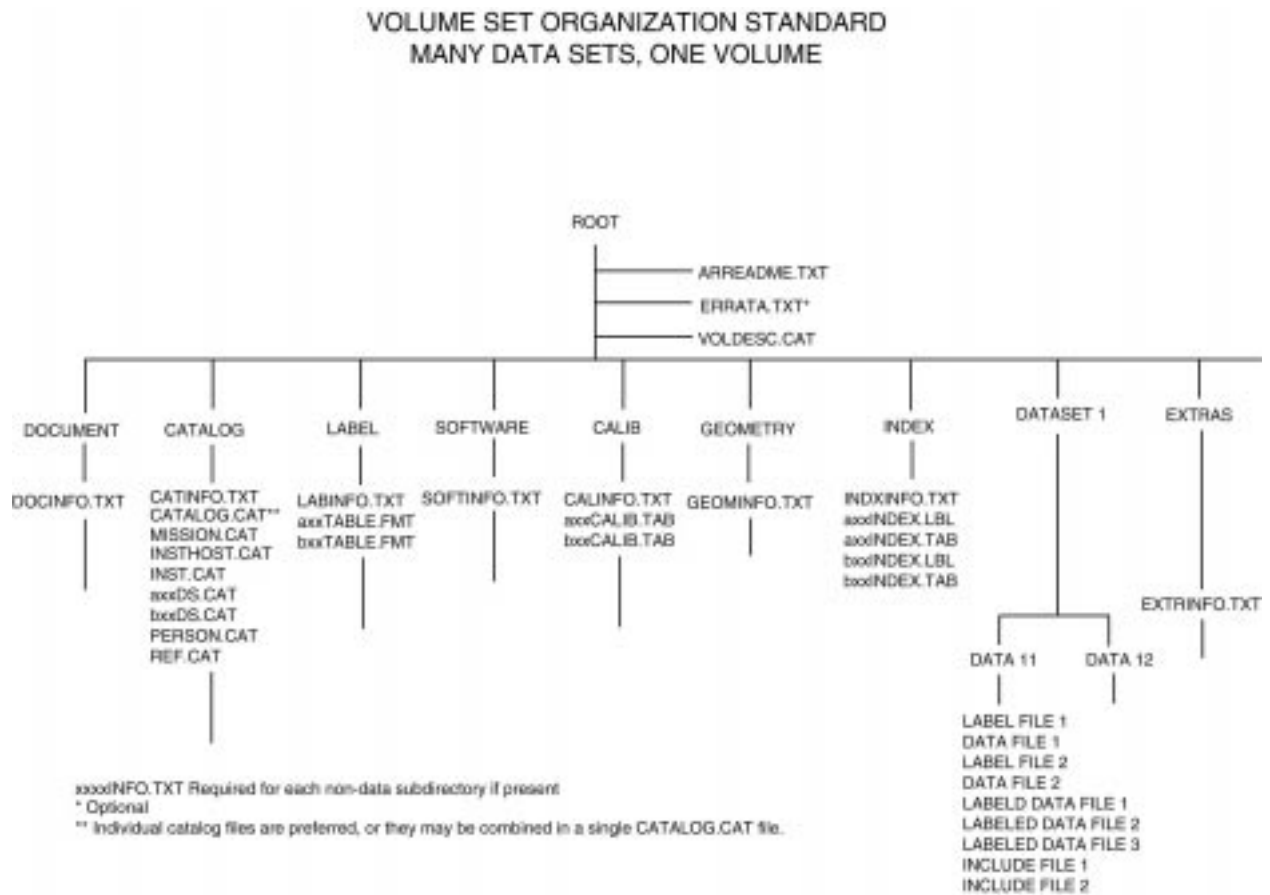


Figure 19.3 Volume Set Organization Standard - Many Data Sets, One Volume

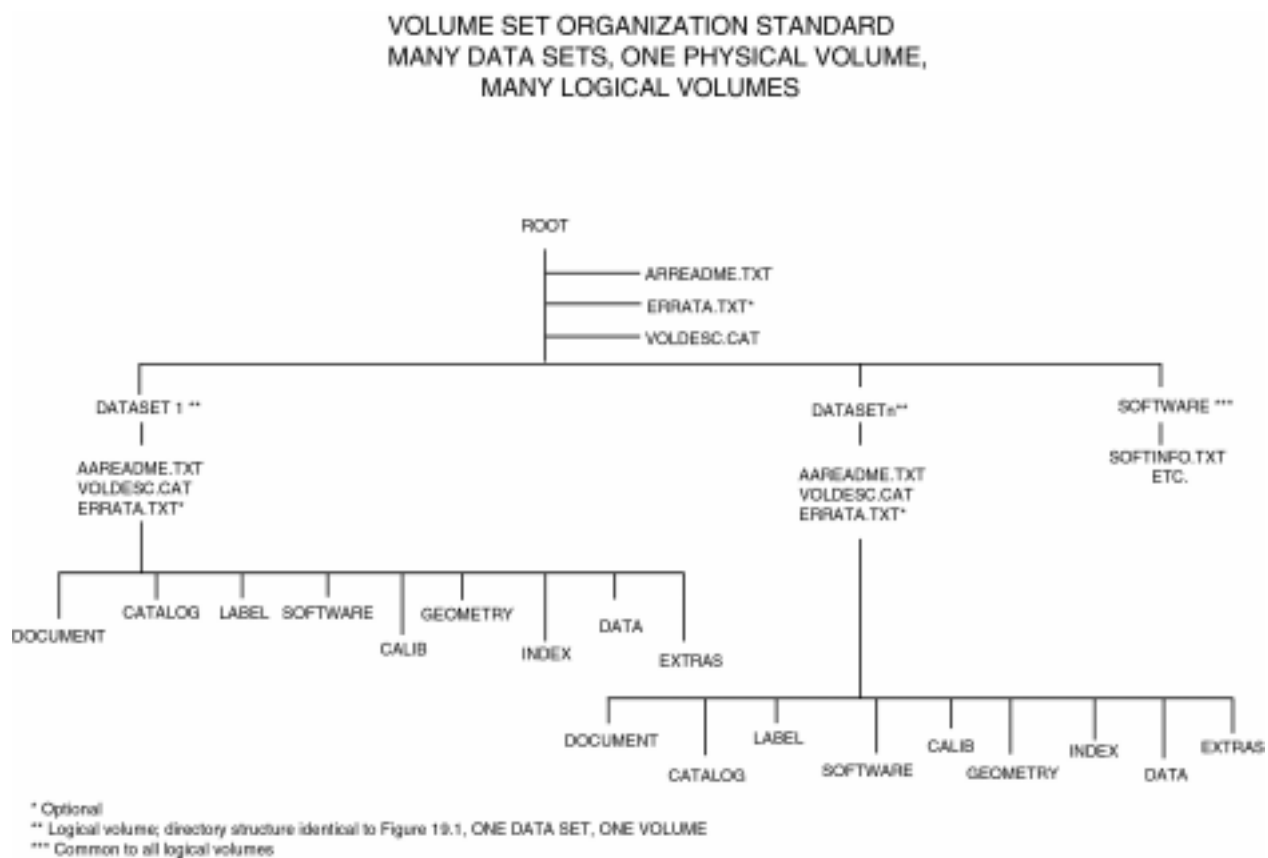


Figure 19.4 Volume Set Organization Standard - Many Data Sets, One Physical Volume, Many Logical Volumes

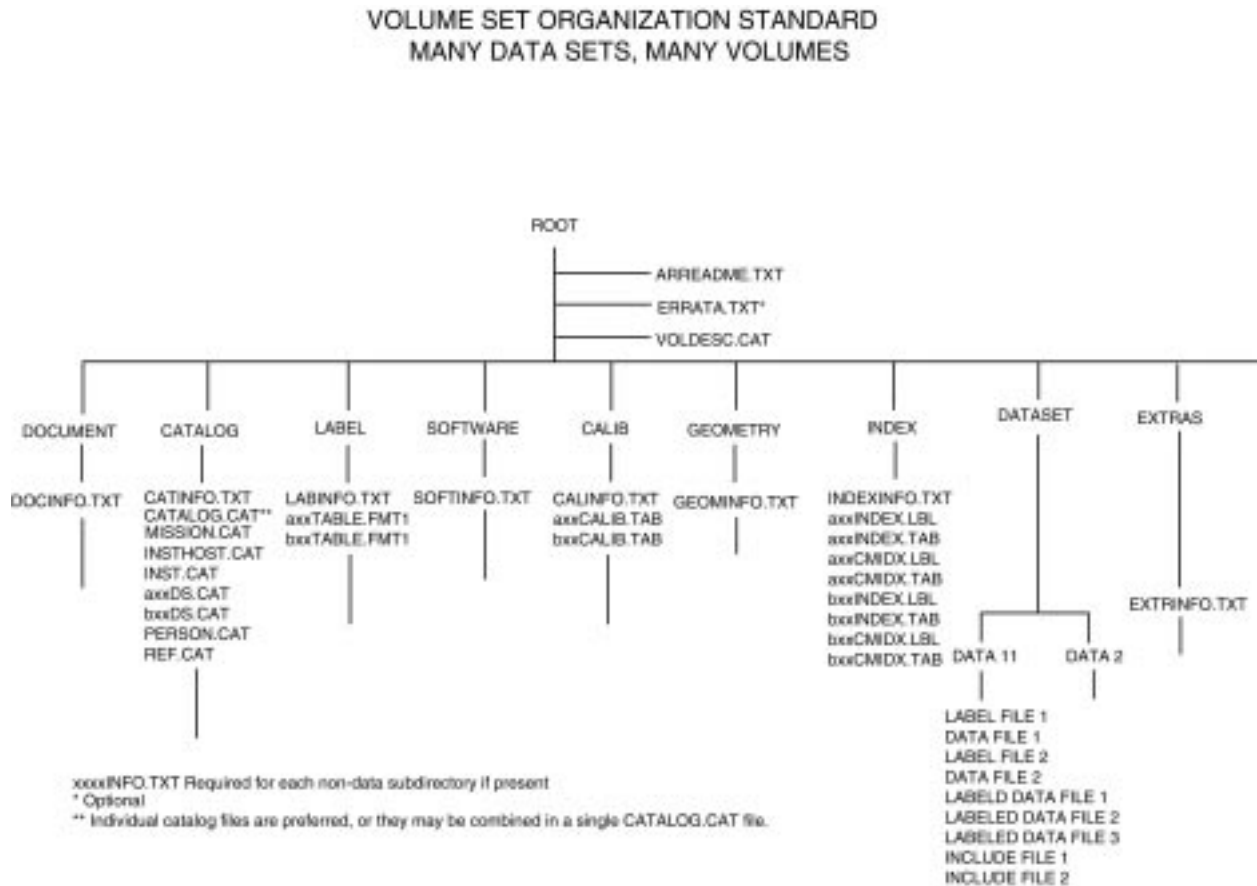


Figure 19.5 Volume Set Organization Standard - Many Data Sets, Many Volumes

19.2 Volume Organization Guidelines

The PDS recommends that directory structures be simple, path names short, and directory and file names constructed in a logical manner. When determining the number of files to be stored in each subdirectory, data preparers should keep in mind that most users rely on visual inspection to glean the contents of a directory or confirm that a disk is intact. Note that some older operating systems will “crash” when encountering a directory containing more than 128 files. Note also that device load time can be directly dependent on the number of files in a directory, making large directories inconvenient for large numbers of users. The typical practical limit for these purposes is on the order of 100 files per directory. As a further convenience to users, PDS recommends that empty subdirectories be omitted entirely.

19.3 Description of Directory Contents and Organization

The root directory is the top-level directory of a volume. The following sections describe the contents of the root directory, followed by the contents of the required subdirectories (in alphabetical order), and finally the contents of the optional directories (in alphabetical order).

19.3.1 ROOT Directory Files

AAREADME.TXT

Required

This file contains an overview of the contents and organization of the associated volume, general instructions for its use, and contact information. The name has been chosen so that it will be listed first in an alphabetical directory listing. See Appendix D for an example of an AAREADME.TXT file.

VOLDESC.CAT

Required

This file contains the VOLUME object, which gives a high-level description of the contents of the volume.

ERRATA.TXT

Optional

This file identifies and describes errors and/or anomalies found in the current volume, and possibly in previous volumes of a set. When a volume contains known errors they must be documented in this file.

VOLDESC.SFD

Obsolete

This file is identified here only for backward compatibility with previous versions of the PDS standards. It is not to be used in current archive products.

This file contains the SFDU reference object structure that aggregates the separate file contents of the volume into an SFDU. The reference object itself is expressed in ODL. This file should only be included if the data products are packaged as SFDUs. (Note the “.SFD” file extension is a reserved file extension in the CCSDS SFDU standard indicating the file contains a valid SFDU.)

19.3.2 CATALOG Subdirectory (Required)

This subdirectory contains the catalog object files (for the mission, instrument, data sets, etc.) for the entire volume. When several logical volumes are present on a single physical volume, each logical volume should have its own CATALOG subdirectory.

CATINFO.TXT

Required

This file identifies and describes the function of each file in the CATALOG subdirectory.

CATALOG.CAT

Optional

In most cases, the individual catalog objects are in separate files, one for each object. On some older archive volumes, however, all catalog objects were collected into a single file called CATALOG.CAT.

PDS Methodology for Supplying Catalog Objects

The preferred method for supplying catalog objects is as separate files for each catalog object, since this facilitates the review, verification and archiving process. In Figure 19.5, for example, the files *axxxxDS.CAT* and *bxxxxDS.CAT* represent two separate files each containing single data set catalog objects (descriptive information about the data set) for data sets *a* and *b* respectively. See the *File Specification and Naming* chapter in this document for the file naming rules; see also *Appendix A* for the required contents of the catalog object.

When catalog objects are organized in separate files or sets of files, pointer expressions shall be constructed according to the following table. Under "File Name", the first line shows the file name to be used if a single catalog file is present on the volume for the particular type of catalog object named. The second shows the syntax and file name convention to be followed if multiple catalog files are present for the named object.

Catalog Pointer Name	File Name
^DATA_SET_CATALOG	= "DATASET.CAT"
	= {"xxxxxDS.CAT", "yyyyyyDS.CAT"}
^DATA_SET_COLLECTION_CATALOG	= "DSCOLL.CAT"
	= {"xxxxxDSC.CAT", "yyyyyDSC.CAT"}
^DATA_SET_MAP_PROJECTION_CATALOG	= "DSMAP.CAT"
	= {"xxxDSMAP.CAT", "yyyDSMAP.CAT"}
^INSTRUMENT_CATALOG	= "INST.CAT"
	= {"xxxxINST.CAT", "yyyyINST.CAT"}
^INSTRUMENT_HOST_CATALOG	= "INSTHOST.CAT"
	= {"xxxxHOST.CAT", "yyyyHOST.CAT"}
^MISSION_CATALOG	= "MISSION.CAT"
	= {"xxxxxMSN.CAT", "yyyyyMSN.CAT"}
^PERSONNEL_CATALOG	= "PERSON.CAT"
	= {"xxxxPERS.CAT", "yyyyPERS.CAT"}
^REFERENCE_CATALOG	= "REF.CAT"
	= {"xxxxREF.CAT", "yyyyyREF.CAT"}
^SOFTWARE_CATALOG	= "SWINV.CAT"
	= {"xxxSWINV.CAT", "yyySWINV.CAT"}
^TARGET_CATALOG	= "TARGET.CAT"
	= {"xxxTGT.CAT", "yyyTGT.CAT"}

19.3.3 Data Subdirectory (Required)

The DATA subdirectory is used to unclutter the root directory of a volume by providing a single entry point to multiple data subdirectories. These directories contain the data product files. The directories are organized and named according to the standards in Chapter 8, *Directory Types and Naming*, in this document. Subdirectories may be nested up to eight levels deep on a physical volume.

Note: Data subdirectories are not present at the root level of a physical volume when multiple logical volumes (case 3b of section 19-1) are present. Instead, they appear under the root directories of each logical volume.

Data Files

A data file contains a data object, which is a grouping of data resulting from a scientific observation (such as an image or table) and representing the measured instrument parameters.

Label Files

A label file contains a detached PDS label that identifies, describes, and defines the structure of the data objects. The associated data objects are contained in an accompanying data file. The label file must have the same base name as the associated data file, with an extension of “.LBL”.

Labeled Data Files

PDS labels may be attached directly to the data they describe. In this case the PDS label comes first and the data begin immediately following the end of the label. When detached labels are used, no “.LBL” files will be present in the data directories. See the *Data Products* and *Data Product Labels* chapters in this manual for details.

19.3.4 INDEX Subdirectory (Required)

This directory contains the indices for all data products on the volume.

Note: If the physical volume is organized as several logical volumes (case 3b of section 19-1), there will generally not be an INDEX subdirectory at the root of the physical volume. Instead there will be individual INDEX subdirectories at the root of each logical volume.

INDXINFO.TXT

Required

This file identifies and describes the function of each file in the INDEX subdirectory. This description should include at least:

- 1) A description of the structure and contents of each index table in this subdirectory
- 2) Usage notes

For an example of the INDXINFO.TXT file, see Appendix D, Section D.2.

INDEX.LBL

Required

This is the PDS label for the volume index file, INDEX.TAB. The INDEX_TABLE specific object should be used to identify and describe the columns of the index table. See *Appendix A* for an example. Although INDEX.LBL is the preferred name for this file, the name *axx*INDEX.LBL may also be used (with *axx* replaced by an appropriate mnemonic).

Note: The PDS recommends detached labels for index tables. If an attached label is used, this file is omitted.

INDEX.TAB**Required**

This file contains the volume index in tabular format (i.e., the INDEX_TABLE specific object is used to identify and describe the data stored on an archive volume). Only data product label files (i.e., not the data files) are included in an index table. In rare cases, however, ancillary files are also included. Although INDEX.TAB is the preferred name for this file, the name *axx*INDEX.TAB may also be used, with *axx* replaced by an appropriate mnemonic.

Note that the *axx* prefix is neither required nor recommended. Data producers may use a prefix to distinguish two or more files by data set, instrument, or other criteria. The data producer should replace the generic prefixes shown here with a suitable mnemonic.

The following files are recommended for multi-volume sets:

CUMINDEX.LBL**Optional**

This file contains the cumulative volume set index in tabular format (i.e., the INDEX_TABLE specific object is used to identify and describe the data stored on each archive volume). Only data product label files (i.e., not the data files) are included in an index table. In rare cases, however, ancillary files may be included. Although CUMINDEX.LBL is the preferred name for this file, the name *axx*CMIDX.LBL may also be used, with *axx* replaced by an appropriate mnemonic.

PDS recommends the use of detached labels for index tables. If an attached label is used, this file is omitted.

CUMINDEX.TAB**Optional**

This file contains the cumulative volume set index in a tabular format. Normally only data files are included in a cumulative index table. In some cases, however, ancillary files may be included. Although CUMINDEX.TAB is the preferred name for this file, the name *axx*CMIDX.TAB may also be used, with *axx* replaced by an appropriate mnemonic.

19.3.5 CALIBration Subdirectory (Optional)

This directory contains the calibration files used in the processing of the raw data or needed to use the data products on the volume. Note that “CALIB” is only a recommended name - a different directory name may be used if appropriate.

CALINFO.TXT**Required**

This file identifies and describes the function of each file in the CALIB subdirectory.

Calibration Files**Required**

In Figures 19.3 and 19.5, the files *axx*CALIB.TAB and *bxx*CALIB.TAB represent sample files. The *axx* and *bxx* prefixes indicate that the calibration files for different data sets (*a* and *b*) may be combined in the same CALIB subdirectory.

Note that the *axx* and *bxx* prefixes in the sample names are neither required nor recommended. Data producers may use them to distinguish two or more files (by data set, instrument, or other criteria). Also, in this case the “CALIB” file name is not required. It is used in the figures to differentiate calibration files from observational data files. The data producer should replace the generic file names shown here by suitably mnemonic names.

19.3.6 DOCUMENT Subdirectory (Optional)

This directory contains the files that provide documentation and supplementary and ancillary information to assist in understanding and using the data products on the volume. The documentation may describe the mission, spacecraft, instrument, and data set(s). It may include references to science papers published elsewhere as well as entire papers republished on the volume.

DOCINFO.TXT**Required**

This file identifies and describes the function of each file in the DOCUMENT subdirectory.

VOLINFO.TXT**Optional**

This file describes the attributes and contents of the volume. This file is sometimes included in addition to the catalog files in the CATALOG subdirectory to provide the same information in an alternate format.

Note: In rare cases, the data engineer may allow the data preparer to place all the corresponding catalog object descriptions in the VOLINFO.TXT file of the DOCUMENT subdirectory in lieu of separate files in the CATALOG subdirectory. Regardless of which method is used, the descriptions themselves must always be supplied.

19.3.7 EXTRAS Subdirectory (Optional)

The EXTRAS directory is the designated area for housing additional elements provided by data preparers beyond the scope of the PDS archive requirements. Examples include HTML-based disk navigators, educational and public interest aids, and other useful but nonessential items. The PDS places no restrictions on the contents and organization of this subdirectory other than conformance to ISO-9660/UDF standards.

EXTRINFO.TXT**Required**

This file identifies and describes the function of each file in the EXTRAS subdirectory. This description should include at least the following:

1. A description of the structure and contents of each file in the subdirectory
2. Usage notes

19.3.8 GAZETTER Subdirectory (Optional)

This directory contains detailed information about all the named features on a target body (i.e., the gazetteer information) associated with the data sets on the volumes. “Named features” are those the International Astronomical Union (IAU) has named and approved.

GAZINFO.TXT**Required**

This file identifies and describes the function of each file in the GAZETTER subdirectory.

GAZETTER.TXT**Required**

This file contains text describing the structure and contents of the gazetteer table in GAZETTER.TAB.

GAZETTER.LBL**Required**

This file is the PDS label containing a formal description of the structure of the gazetteer table.

GAZETTER.TAB**Required**

This file contains the gazetteer table.

19.3.9 GEOMETRY Subdirectory (Optional)

This directory contains the files (e.g., SEDR file, SPICE kernels, etc.) needed to describe the observation geometry for the data. Note that “GEOMETRY” is only a recommended directory name, another appropriate name may be used.

GEOMINFO.TXT**Required**

This file identifies and describes the function of each file in the GEOMETRY subdirectory.

19.3.10 LABEL Subdirectory (Optional)

This directory contains additional PDS labels and include files that were not packaged with the data products or in the data subdirectories. When multiple logical volumes reside on a single physical volume, the LABEL subdirectories *must* appear below the logical volume root directories. This is because the rules governing pointer resolution preclude a search across logical volumes.

LABINFO.TXT**Required**

This file identifies and describes the function of each file in the LABEL subdirectory.

Include Files**Required**

Include files are files referenced by a pointer in a PDS label. Typically they contain additional metadata or descriptive information. Only files of type LBL, TXT, or FMT (“format”) may be included in the LABEL subdirectory. In Figures 19.1-5, the files *axx*INCLUDE FILE1, *bxx*INCLUDE FILE1 and INCLUDE FILE1 represent sample files of the above types. The *axx* and *bxx* prefixes indicate that the include files for different data sets (*a* and *b*) may be combined in the same LABEL subdirectory.

Note that the *axx* and *bxx* prefixes in the sample names are neither required nor recommended.

Data producers may use them to distinguish two or more files (by data set, instrument, or other criteria). The data producer should replace the generic prefixes shown here by a suitable mnemonic.

19.3.11 SOFTWARE Subdirectory (Optional)

This directory contains the software libraries, utilities, or application programs supplied for accessing or processing the data. It may also include descriptions of processing algorithms. Only public domain software may be included on PDS archive volumes.

Two subdirectory structures are available for organizing the SOFTWARE directory: platform-based and application-based. Platform-based is the recommended method for general archives and is described below. For an example of application-based organization see the example for SOFTINFO.TXT in *Appendix D* of this document, and the NAIF directory structure in *Appendix E*. See Section 11.3 for information about packaging software for inclusion in an archive product.

SOFTINFO.TXT

Required

This file identifies and describes the function of each file in the SOFTWARE subdirectory.

SRC Subdirectory

Optional

There can be a global SRC directory under the SOFTWARE directory if there is source code applicable to all platforms. For example, application-programming languages such as IDL are relatively platform independent and would be placed in a global SRC directory. Note that in the example below, there is both a global source directory as well as source directories at the lower levels.

DOC Subdirectory

Optional

This directory contains documentation for the software in the parallel SRC directory.

LIB Subdirectory

Optional

This directory contains libraries applicable to all platforms.

Hardware Platform and Operating System/Environment Subdirectories

Optional

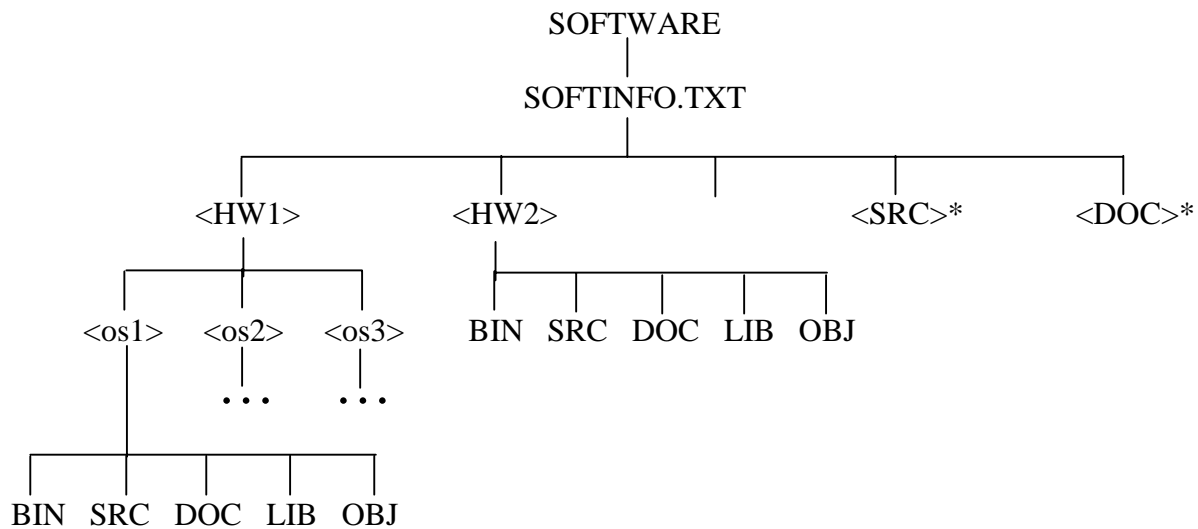
If only global source code is being provided on the volume, no further organization is required. If platform- or environment- specific software is being provided, the structure in Figure 19.6 should be followed. Specifically:

1. The hardware platform and the operating system/environment must be explicitly stated. If more than one operating system/environment (OS/Env) is supported for a single hardware platform, each should have its own subdirectory under the hardware directory.

If there is only one, then that subdirectory can be promoted to the hardware directory level (via naming conventions). In Figure 19.6, several environments are supported for platform HW1, but only one for HW2 – thus the difference in subdirectory structures.

2. The next directory level contains BIN, SRC, DOC, LIB and OBJ. If any of these are not applicable, it should be left out (i.e., empty directories should be omitted).
3. Following are examples of subdirectory names for both multiple and single OS/Env per platform. (This list is provided for illustration only. It is not meant to be exhaustive.)

Multiple	Single
<i>PC</i>	
DOS	PCDOS
WIN	PCWIN
WINNT	PCWINNT
OS2	PCOS2
<i>MAC</i>	
SYS7	MACSYS7
AUX	MACAUX
<i>SUN</i>	
SUNOS	SUNOS
SOLAR	SUNSOLAR
<i>VAX</i>	
VMS	VAXVMS
ULTRX	VAXULTRX
<i>SGI</i>	
IRX4	SGIIRX4
IRX5	SGIIRX5



* **NOTE:** INFO.TXT files under SOFTWARE subdirectories are optional (e.g., PCINFO.TXT, MACINFO.TXT, VAXINFO.TXT, SUNINFO.TXT, etc.).

Figure 19.6 – Platform-based SOFTWARE Subdirectory Structure

19.4 Volume Naming

Volume names must be no more than 60 characters in length and in upper case. They should describe the contents of the volume in terms that a human user can understand. In most cases the volume name is more specific than the volume set name. For example, the volume name for the first volume in the *VOYAGER IMAGES OF URANUS* volume set is “VOLUME 1: COMPRESSED IMAGES 24476.54 - 26439.58.”

19.4.1 Volume ID

Many types of media and the machines that read them place a limit on the length of the volume ID. Therefore, although the complete volume set ID should be placed on the outside label of the volume, a shorter version is actually used when the volume is recorded. PDS has adopted a limit of 12 characters for these terse volume identifiers. This volume ID consists of the last two components of the volume set ID, with the “X” wildcard values replaced by the sequence number associated with the particular volume (see the *Volume Set ID* Standard below). This ID must always be unique for PDS data volumes. The volume ID must be in upper case.

Examples:

VG_0002	Volume 2 of the Voyager set
MG_0001	The first volume of the Magellan set
VGRS_0001	A potential Voyager Radio Science collection

If a volume is redone because of errors in the initial production the volume ID should remain the same and the `VOLUME_VERSION_ID` incremented. This parameter is contained in the `VOLDESC.CAT` file on the volume. The version ID should also be placed on the external volume label as “Version *n*” where *n* indicates the revision number. A revision number greater than one indicates that the original volume should be replaced with the new version.

19.5 Volume Set Naming

The volume set name provides the full, formal name of a group of data volumes containing one or a collection of related data sets. Volume set names may be at most 60 characters in length and must be in upper case. Volume sets are normally considered a single orderable entity. For example, the volume series `MISSION TO VENUS` consists of the following volume sets:

MAGELLAN: THE MOSAIC IMAGE DATA RECORD

MAGELLAN: THE ALTIMETRY AND RADIOMETRY DATA RECORD

MAGELLAN: THE GLOBAL ALTIMETRY AND RADIOMETRY DATA RECORD

PRE-MAGELLAN RADAR AND GRAVITY DATA SET COLLECTION

In certain cases, the volume set name can be the same as the volume name, e.g., when the volume set consists of only one volume.

19.5.1 Volume Set ID

A *volume set* is a series of archive volumes that are closely related. In general, the volumes of a set will be distributed and used together. Each volume within the set must have a `VOLUME_ID` that is unique across the PDS archive. The volume set is identified by a `VOLUME_SET_ID` of up to 60 characters incorporating the range of constituent `VOLUME_ID`s. `VOLUME_SET_ID`s must be in upper case, and are composed by concatenating the following fields, separated by underscores, using abbreviations if necessary:

1. The country of origin (abbreviated)
2. The government branch
3. The discipline within the branch that is producing the volumes
4. A campaign, mission or spacecraft identifier (2 characters) followed by an optional 2-character instrument or product identifier
5. A 4-digit sequence identifier: The first digit(s) represent the volume set; the remaining digits contain “X”, representing the range of volumes in the set. Up to four “X” characters may be used.

Example

`USA_NASA_PDS_GO_10XX` could be the volume set ID for the Galileo EDR volume set, since there are less than 100 volumes (since the `XX` placeholder accommodates the range 01 - 99 only). Volume IDs for volumes in the set would then be `GO_1001`, `GO_1002`, etc.

Note: Because of the uniqueness constraint, data preparers should consult with their PDS data engineer when it comes time to formulate new VOLUME_ID and VOLUME_SET_ID values.

Volume Set IDs Prior to PDS Version 3.2

Prior to version 3.2, the 4-digit sequence identifier (item 5 above) did not include the “X” wildcards. Instead, the last digits represented the volume. For example, on Magellan, a volume set ID “USA_NASA_JPL_MG_0001” was used *only* for the volume with the volume ID “MG_0001”. Subsequent volumes in the same set had volume set IDs that differed in the final field. When a set of volumes was to be distributed as one logical unit, the volume set ID included the range of volume IDs.

Example

USA_NASA_PDS_VG_0001_TO_VG_0003 for the three volumes that comprise the Voyager Uranus volume set.

19.6 Logical Volume Naming

Logical volumes retain the volume and volume set naming used at the physical volume level. For further information, see the “Volume Object” in *Appendix A* of this document.

19.7 Exceptions to This Standard

In rare cases volume IDs are subject to restrictions imposed by specific hardware or software environments. Also, volumes made in the past may have IDs that do not meet this standard and there may be compelling reasons for keeping the same volume ID when making a new copy of the data. All new data sets, however, must adhere to this standard wherever possible.